Plant Gene 5 (2016) 100–108

Contents lists available at ScienceDirect

Plant Gene

journal homepage: www.elsevier.com/locate/plantgene

Molecular cloning, gene expression profiling and in silico sequence analysis of vitamin E biosynthetic genes from the oil palm

Kong Sze Ling [a], Siti Nor Akmar Abdullah [a,b,*], Ho Chai Ling [c], Mohd Din Amiruddin [d]

[a] Laboratory of Plantation Crops, Institute of Tropical Agriculture, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia
[b] Department of Agriculture Technology, Faculty of Agriculture, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia
[c] Department of Cell and Molecular Biology, Faculty of Biotechnology and Biomolecular Sciences, University Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia
[d] Advanced Biotechnology and Breeding Centre, Malaysian Palm Oil Board, No. 6, Persiaran Institusi, B. B. Bangi, 43000 Kajang, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 30 July 2015
Received in revised form 6 January 2016
Accepted 18 January 2016
Available online 26 January 2016

Keywords:
Oil palm
Homogentisate phytyltransferase (HPT)
Homogentisate geranylgeranyl transferase (HGGT)
Degenerate primer
qPCR
Promoter analysis

ABSTRACT

Homogentisate geranylgeranyl transferase (HGGT) and homogentisate phytyltransferase (HPT) are the two key enzymes involved in condensation of homogentisic acid (HGA) with a prenyldiphosphate to produce tocotrienols and tocopherols in plants, respectively. The partial cDNAs encoding HGGT and HPT enzymes were successfully isolated from the two oil palm species, *Elaeis guineensis* and *Elaeis oleifera* by PCR amplification using degenerate primers. Subsequently, full length cDNA sequences were completed by rapid amplification of cDNA ends (RACE) and further annotated using various bioinformatics tools. The analysis revealed the presence of an UbiA prenyltransferase conserved domain in all four deduced amino acid sequences and suggested that oil palm HGGT and HPT are more evolutionarily related to their counterparts from other monocot plant species. Quantitative gene expression analysis was carried out to elucidate the transcript profiles of the oil palm HGGT and HPT in different oil palm tissues and at different developmental stages of the mesocarp. The HPT was constitutively expressed in all analyzed tissues except in 15 w.a.a kernel whereas oil palm HGGT showed preferential expression in mesocarp and kernel tissues. However, HPT was highly expressed at the fruit ripening stage of 17 w.a.a mesocarp when active oil deposition occurs. Genome-walking PCR successfully amplified the promoter regions of HGGT and HPT from *E. guineensis*. Computational analysis using PlantCare and PLACE databases revealed several *cis*-regulatory elements including phytohormone-responsive, light-responsive and abiotic factor-responsive elements which may be involved in coordinating expression of both genes. Taken together, this study provides useful information about important features of the cDNA and promoter sequences as well as an insight into the transcriptional regulation of these key vitamin E genes for future genetic improvement efforts.

1. Introduction

Tocochromanols, commonly known as vitamin E, which are exclusively synthesized by photosynthetic organisms, are essential fat soluble nutrients in the human diet (Dörmann, 2003). The eight structurally related forms of tocopherols and tocotrienols (α-, β-, γ-, and δ-tocopherols and α-, β-, γ-, and δ-tocotrienols) together constitute the content of vitamin E. Tocotrienols structurally differ from tocopherols by the presence of three *trans* double bonds in the hydrocarbon tail while the aliphatic tail of tocopherols is fully saturated (Kamal-Eldin and Appelqvist, 1996). Vitamin E is well accepted as nature's most effective chain-breaking antioxidant that prevents the degradation of polyunsaturated fatty acids in membrane by reactive oxygen species (Dörmann, 2003). It was suggested that tocotrienols possess powerful neuroprotective, antioxidant, anti-cancer and cholesterol lowering properties which are not often exhibited by tocopherols (Sen et al., 2006). The pharmaceutical properties demonstrated by tocotrienols would make tocotrienol-accumulating crops potentially very beneficial for human health.

Tocochromanols also play important roles in plants, such as protection of chloroplasts from photooxidative damage (Munné-Bosch and Alegre, 2002). As a strong antioxidant, vitamin E is beneficial in

Abbreviations: BLAST, basic local alignment search tool; bp, base pair; cDNA, complementary deoxyribonucleic acid; Ct, threshold cycle; CTAB, hexadecyl (or cetyl) trimethyl ammonium bromide; Da, Dalton; DNA, deoxyribonucleic acid; DNase I, deoxyribonuclease I; dNTP, deoxynucleoside triphosphate; GA, gibberellin; GGDP, geranylgeranyldiphosphate; GSP, gene-specific primer; HGA, homogentisic acid; HGGT, homogentisate geranylgeranyl transferase; HPT, homogentisate phytyltransferase; kb, kilobase; min, minutes; mRNA, messenger ribonucleic acid; NCBI, National Center for Biotechnology Information; ORF, open reading frame; PCR, polymerase chain reaction; PDP, phytyldiphosphate; pI, isoelectric point; PrDP, prenyldiphosphate; RACE, Rapid Amplification of cDNA End; RNA, ribonucleic acid; RT-PCR, reverse transcription PCR; sec, seconds; TSS, transcription start site; UTR, untranslated region; w.a.a, week after anthesis.

* Corresponding author at: Laboratory of Plantation Crops, Institute of Tropical Agriculture, Universiti Putra Malaysia, 43400, UPM, Serdang, Selangor, Malaysia.
E-mail address: snaa@upm.edu.my (S.N.A. Abdullah).

http://dx.doi.org/10.1016/j.plgene.2016.01.003

maintaining oxidative stability of vegetable oils and also in enhancing nutritional value of crop plants for human diets and livestock feeds (Hunter and Cahoon, 2007). Studies have shown that tocopherols occur ubiquitously in various plant tissues especially in leaves and seeds of most dicots and are also found in photosynthetic microbes such as *Synechocystis*. Tocotrienols, in contrast, are rarely found in vegetative tissues of plants but are the exclusive form of vitamin E present in the seed endosperm of most monocots, including agronomically important cereal grains such as wheat, rice and barley (Kamal-Eldin and Appelqvist, 1996). Soybean oil and palm oil are the main sources for commercially produced tocopherols and tocotrienols, respectively. These lipid-soluble products are found as vegetable oils components (Hunter and Cahoon, 2007). Palm oil is unique among the vegetable oils because it contains high level of tocotrienols (α, γ, and δ), which together account for about 70% of the total vitamin E content. Tocochromanols are found in the crude palm oil at levels of 600–1000 ppm (Choo et al., 2004). Palm vitamin E has been extensively studied for its nutritional and health properties, attributed largely to its high tocotrienol content (Sen et al., 2006).

The biosynthesis pathway of tocochromanols can be divided into four stages starting from the formation of the homogentisic acid head group. HGA is the first substrate needed in the biosynthesis of vitamin E by supplying the aromatic ring of the chromanol head group and it is derived from the shikimate pathway. The second required substrate is a prenyldiphosphate (PrDP), either phytyldiphosphate (PDP) or geranylgeranyldiphosphate (GGDP). These prenyl groups are supplied by the non-mevalonate pathway. The second stage is the transfer of a prenyl group to HGA which is the branch point for commitment to either tocopherols or tocotrienols and is a potential key regulatory step. Homogentisate geranylgeranyl transferase (HGGT) is generally believed to be the key enzyme that catalyzes the condensation of HGA and GGDP to generate tocotrienols due to the structural similarity of the tocotrienols side chain and GGDP. Meanwhile the committed step in the formation of tocopherols is the condensation of HGA with PDP through an enzyme designated as homogentisate phytyltransferase (HPT). The last two stages are the subsequent ring cyclization and methylation reactions to generate all eight forms of tocochromanol (Hunter and Cahoon, 2007; Collakova and DellaPenna, 2001; Bramley et al., 2000).

HPT was first identified from *Arabidopsis* and *Synechocystis* sp. PCC 6803 and it was further demonstrated that the level of seed tocopherols could be elevated up to 2-fold by seed-specific expression of the *Arabidopsis HPT1* in *Arabidopsis* (Collakova and DellaPenna, 2001; Savidge et al., 2002). In 2003, Cahoon and coworkers isolated the gene encoding HGGT from three monocot species (*Oryza* sp., *Hordeum* sp. and *Triticum* sp.) and over-expressed it in corn seeds driven by a strong embryo-specific promoter. This was resulted in the increment of tocotrienol content by 20-fold and also enhanced the vitamin E content by 6-fold. Oil palm may offer great advantages compared to other plants for genetic manipulation of vitamin E. However, the knowledge on biosynthesis pathway of vitamin E in oil palm as one of the basic prerequisites for genetic manipulation is quite limited. This will definitely become an impediment to improve oil palm vitamin E content through genetic engineering, development of molecular markers for cross species breeding as well as the other biotechnological approaches. The present study reports on characterization of the cDNA and promoter sequences of the HGGT and HPT in two important oil palm species. Quantitative gene expression analysis in various oil palm tissues was also performed to obtain an insight into the transcriptional regulation of these key vitamin E genes.

2. Materials and methods

2.1. Plant materials

The mesocarp and kernel tissues of two oil palm species, *Elaeis guineensis* and *Elaeis oleifera* were harvested at various weeks after anthesis (7-, 10-, 12-, 15-, 17- and 19-w.a.a) and sliced into small pieces, snap frozen in liquid nitrogen and stored at −80 °C. All fruit bunches were obtained from tagged fruits in the same field. Unfolded spear leaves and young roots from *E. guineensis* seedlings were harvested and frozen at −80 °C.

2.2. Total RNA isolation and cDNA synthesis

Total RNA extraction was carried out using the modified method of Prescott and Martin (1987) and all of the extracted RNAs were treated with DNaseI (Fermentas, USA). The amount of total RNA was quantified using a Nanodrop spectrophotometer (ND-1000, Thermo Scientific, USA). The integrity of the total RNA was checked using 1% (*w*/*v*) agarose gel electrophoresis and separated at 75 V in 1× TAE buffer. Mesocarp Poly(A)$^+$ RNAs of both oil palm species at 17 w.a.a were purified using the Absolutely mRNA™ Purification Kit (Stratagene, CA, USA) and used as the template to synthesize the cDNA pool for gene isolation. First-strand cDNA was synthesized using Superscript III First-Strand Synthesis Kit (Invitrogen, USA) and 150 ng of each mRNA template in a final volume of 20 μl.

2.3. Isolation of the full length HGGT and HPT cDNAs

RT-PCR amplification was conducted using Advantage cDNA polymerase mix (Clontech, USA) and a pair of sense and antisense degenerate oligonucleotides which designed based on the conserved nucleotide region of known monocot HGGT and HPT protein sequences (Table 1). A similar forward degenerate primer (F5-2) was used to amplify both genes in oil palm. The PCR reaction was performed in a T-Professional Basic Thermal Cycler (Gradient) (Biometra, Germany) using an initial denaturation step at 94 °C for 5 min, followed by 35 cycles at 94 °C for 30 s, 58 °C for 30 s and 72 °C for 1 min and final extension at 72 °C for 5 min. The desired PCR fragment was cloned into pGEM-T Easy vector (Promega, USA) and sequenced. A homology search was conducted through NCBI database using BLASTX and BLASTP algorithms (http://www.ncbi.nlm.nih.gov/). Gene-specific primers (GSPs) (Table 1) were designed based on the isolated partial cDNAs for amplification of 5′- and 3′-ends of the oil palm HGGT and HPT genes by RACE PCR. The 5′- and 3′-RACE ready cDNA synthesis was carried out using SMARTer™ RACE cDNA Amplification Kit (Clontech, USA). Both 5′- and 3′-sequences were combined to produce the full length gene sequences. For long-distance PCR (LD-PCR), new gene specific primers were designed based on the identified 5′- and 3′-end sequences covering the open reading frame (ORF) of the oil palm HGGT and HPT genes. The primers were designated as follows: (1) *E. guineensis*; HGGT forward primer EEF1 (5′-GGTTGCAACGCATCATTAAGATATTG-3′), HGGT reverse primer EER1 (5′-GGAGAATTTTCAAAAGTCTATGATGATG-3′). HPT forward primer EGHPTF (5′-GGTTCTCTTACCAGGTCCCTGCAAGATCG-3′), HPT reverse primer EOHPTR (5′-CATAATGGATGATCGAAGTGGCACAG-3′); (2) *E. oleifera*; HGGT forward primer EEF2 (5′-GCTTCAAAATGACC TGCTTTTTATGG-3′), HGGT reverse primer EER1, HPT forward primer EOHPTF (5′-GGTCTCTTACCAGATCCCTGCAAGATCG-3′) and HPT reverse primer as EOHPTR.

2.4. Full length cDNA sequence analysis

Initially, ORFs for all four cDNAs were predicted using the ORF Finder program (http://www.ncbi.nlm.nih.gov) and subjected to BLASTX and BLASTP analyses. The sequence alignment and characterization of the deduced proteins were carried out using different tools available in SDSC Biology Workbench (http://seqtool.sdsc.edu). Multiple sequence alignment was done using ClustalW pairwise alignment algorithm. To find out general features of the cDNA sequence including amino acid composition and isoelectric point (pI), the analysis was performed using ExPASy Proteomics tools (http://cn.expasy.org/tools/protscale.html). Physical and chemical characteristics of all deduced amino acid

Table 1
List of primers used in gene isolation and cloning reactions. The degenerate genetic code was according to Sambrook et al. (2001).

Reaction	Target gene	Primer name	Nucleotide sequence (5′–3′)
Degenerate PCR	HGGT	F5-2 (Forward)	ATGAAYATYTAYGTGGTNGGNYTGAATCA
		R3-1 (Reverse)	ATACTCHGCRTARAABAGCTTCCAWATGAACAT
	HPT	R3-2 (Reverse)	GGDATATCCTTGAATARDGCTATRAC
5′-RACE	*EgHGGT* *EoHGGT*	GR2	GGACTTGGACTGCAACCCCATTGT
3′-RACE	*EgHGGT*	GF2	AGATATGTACTAGGAAGACCAACAGTG
	EoHGGT	EO3F	GTTGTTGTTAACTGCATATGCTACAGC
5′-RACE	*EgHPT* *EoHPT*	PR1	TCTAGCACGATTCCAGAGGACCG
3′-RACE	*EgHPT* *EoHPT*	PF1	AGCATCTGGGGAATATTCTGTCAC
Genome-walking PCR	*EgHGGT*	HGW1	GAGACTGAAACAGCCTGGTTGGACAGC
		HGW1.5	ACAGAGGAGGAGCACGGCGAGGAGAGG
	EgHPT	HPW2.5	GGAGAAGAAAGAGAATCCATCAAACTTGGC
		HPW1.5	CGATCTTGCAGGGACCTGGTAAGAGAACC
LD-PCR	p*EgHGGT*	GF1	CATTTTTATGTATTTGTGGCAAAAGGG
		GR2	AACTCTTCTTTTGCCATTTGTACAAAGC
	p*EgHPT*	PF4	CTAATCCATTTGCAGTTGTATCGG
		PR1	GAGACTGTTTTCCTTCCGTGAAT

sequences were analyzed by ProtParam tool (http://web.expasy.org/protparam/). The signal peptide targeting location of the deduced proteins was predicted using the SignalP method (http://www.cbs.dtu.dk/services/SignalP) and ChloroP program (http://www.cbs.dtu.dk/services/ChloroP/) while the subcellular localization presumption was performed using WoLF PSORT (http://wolfpsort.org/). Protein domain analysis was done using SMART (Simple Modular Architectural Research Tool) database (http://smart.embl-heidelberg.de/). Phylogenetic and molecular evolutionary analyses of the amino acid sequences of HGGT and HPT from different plant species were constructed using the Phylogeny.fr web services (http://www.phylogeny.fr/version2_cgi/simple_phylogeny.cgi) through Neighbor Joining method.

2.5. Gene expression analysis

In order to validate the stably expressed oil palm reference genes for RT-qPCR normalization, gene expression stability of *β-actin* (ACT), *cyclophilin* (CYP) and *tubulin* (TUB) were studied across three sets of oil palm tissues in various developmental stages. The first and second sets included mesocarps of *E. guineensis* and *E. oleifera* collected at 7, 10, 12, 15, 17 and 19 w.a.a, respectively. Whereas the last set consists of 15 w.a.a kernels sampled from *E. guineensis* and *E. oleifera*, spear leaves and young root of *E. guineensis*. The C_q values were collected for all selected tissues of each oil palm species and together with the gene-specific amplification, efficiencies were transformed to relative quantities using the delta Ct method and entered into geNorm (Vandesompele et al., 2002). Only the top two most stable housekeeping genes in each experiment were used for normalization in the subsequent real-time quantitative reactions. Table 2 listed all of the primers designed for the understudied genes. RT-qPCR was performed on iCycler iQ Real-Time PCR System using the iQ™ SYBR® Green Supermix (Bio-Rad, USA). In each reaction, 200 ng of cDNA was used. We were not able to obtain the C_q within 30 cycles for the amounts lower than 200 ng of template cDNA possibly due to the low abundance nature of these genes. The melt curve result still showed a single peak at this high concentration of cDNA template indicating that the specificity of the reaction was retained. Thermal cycling conditions were: 95 °C for 3 min followed by 40 cycles of denaturation at 95 °C for 30 s, annealing at 57 °C for 30 s and extension at 72 °C for 30 s. Finally, a melting curve was generated by increasing temperature starting from 54 °C to 94 °C with increment of 0.5 °C for each cycle. Each gene was analyzed in three technical replicates with one biological replicate as only one fruit bunch was harvested for each developmental stage. Relative expression level was calculated using the geNorm-derived normalization method.

Table 2
List of the specific primers designed from oil palm HGGT, HPT, *β-actin* (ACT), *cyclophilin* (CYP) and *tubulin* (TUB) for the quantitative PCR assays.

Genes	Primer name	Sequence (5′–3′)	Amplicon length (bp)
β-Actin	ACTF	CAAGTCATGTAGGGTTGC	121
	ACTR	ACTAGGCTGGCAAAGTTCAT	
Cyclophilin	CYPF	TACAAGGGGTCGACGTTCCA	142
	CYPR	GCCCGTGTGCTTCTTGATG	
Tubulin	TUBF	GCCTTCGAGCCATCTTCTATGAT	68
	TUBR	AGGCAGCAAGCCATGTACTTAC	
HGGT	HGGTF	GGAGTCCATGCAGGAGTGT	107
	HGGTR	TCTTGACAGCTTGGATTGCT	
HPT	HPTF	GCAGGAACAATTCTCACCAAAC	176
	HPTR	GATCGAAGTGGCACAGTAATACAC	

2.6. Promoter isolation and in silico promoter analysis

Genomic DNA was extracted from *E. guineensis* spear leaf tissues according to the modified CTAB method of Doyle and Doyle (1990). Five blunt-end digested libraries were set up using GenomeWalker™ Universal kit (Clontech, USA) and labeled as DL1 (*Dra* I), DL2 (*Pvu* II), DL3 (*EcoR* V), DL4 (*Stu* I) and DL5 (*Sca* I). Two sets of GSPs consisting primary and nested primers were designed at the region close to the end of the identified 5′ untranslated region (UTR) of HGGT and HPT genes (Table 1). The primary and nested PCR reactions were conducted according to the manufacturer's protocol. The desired PCR fragments were cloned and sequenced. Based on the sequencing result, two pairs of GSPs were designed for the LD-PCR using Advantage® 2 PCR kit (Clontech, USA) (Table 1). The amplification was started with pre-denaturation at 94 °C, 5 min followed by 35 cycles of 94 °C, 30 s; 54 °C, 30 s; 72°, 1 min 30 s and final extension at 72 °C, 5 min. Transcription start site (TSS) of the promoter was determined by alignment with the identified *EgHGGT* and *EgHPT* cDNA sequences, respectively. A promoter motif search was carried out to define putative *cis*-elements in both promoter sequences using PLACE (http://www.dna.affrc.go.jp/PLACE/signalup.html) and PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Each of the regulatory elements

observed was analyzed with previously reported properties of the particular element.

3. Results

3.1. Full length gene isolation and sequence analysis

Due to the lack of enough information available on the oil palm HPT and HGGT, the degenerate primers used in the RT-PCR were synthesized based on the consensus nucleotide sequences generated from the alignment of the monocot HGGT and HPT homologs. This technique has been successfully employed to amplify several orthologs from oil palm (Rasid et al., 2008; Rival et al., 2008). Partial cDNAs of *E. guineensis* HGGT were isolated and the sequence was used to design primers for RACE-PCR. After assembling the 5′ and 3′ regions, a 1853 bp consensus cDNA sequence of *E. guineensis* HGGT was attained and designated as *EgHGGT* (GenBank accession No. KP878511). The longest ORF of 1389 bp encoded for 462 deduced amino acid residues with a predicted molecular mass of 51.6 kDa and a theoretical pI of 9.51. In addition, the ORF was flanked by 194 bp of 5′-UTR and 246 bp of 3′-UTR. For *E. oleifera*, the full length cDNA was 1732 bp in length and designated as *EoHGGT* (GenBank accession No. KP878512). *EoHGGT* contains a 1389 bp ORF flanked by 69 bp of 5-UTR and 246 bp of 3′-UTR. It encodes a predicted polypeptide of 462 aa with a molecular mass of about 51.6 kDa and a deduced isoelectric point of 9.50. Blast analysis revealed that the ORF of both genes was highly identical (99%) to the newly reported HGGT sequence in *E. guineensis*. Besides, they also shared a high similarity with HGGT from other monocot plant species such as *Oryza sativa* (80%) and *Hordeum vulgare* (73%) and subsequently established the identity of the deduced protein sequence corresponding to the predicted ORF. The ClustalW alignments showed that HGGT in both oil palm species shared as high as 99.1% similarity at the protein level and 99.3% similarity at the nucleotide level. There were a total of nine variant nucleotide bases where only two were involving in the exchange from purine to pyrimidine or vice versa. Four of the base changes have led to the changes in amino acid residues which subsequently caused the differences between *EgHGGT* and *EoHGGT* protein sequences.

Similarly after obtaining partial cDNA using degenerate primers, followed by 5′ and 3′RACE and finally end to end PCR, a 1732 bp consensus cDNA sequence of *E. oleifera* HPT was obtained and designated as *EoHPT* (GenBank accession No. KF974530). *EoHPT* contained an ORF of 1203 bp encoding a polypeptide of 400 aa flanked by 120 bp of 5′-UTR and 385 bp of 3′-UTR. For *E. guineensis*, the full length cDNA was named as *EgHPT* (GenBank accession No. KF974529). It was 1762 bp in size containing an ORF of 1203 bp, a 5′-UTR of 123 bp and a 3′-UTR of 411 bp. Both ORFs encoded a polypeptide of 400 aa with a calculated molecular weight of about 44.4 kDa and a theoretical pI of 9.47. Homology search established identity of the deduced amino acid sequences corresponding to the predicted *EgHPT* and *EoHPT* coding regions based on the newly released *HPT1* in *E. guineensis* (99%) and other monocot *HPT*s such as *Zea mays* (70%), *Allium ampeloprasum* (70%) and *Triticum aestivum* (69%). Pairwise comparison of the oil palm *HPT* has unveiled a 99% similarity at the nucleotide level and an even higher similarity (99.5%) at the protein level with only two amino acid differences.

Multiple alignment results indicated that HGGT and HPT within the same oil palm species only share 50% amino acid sequence identity. Similar observation was also described by Cahoon et al. (2003), in which HGGT from rice, wheat and barley shared <50% deduced protein sequence identity with the earlier characterized HPTs from *Arabidopsis thaliana*, soybean, corn and rice. Furthermore, the pairwise alignment also showed that these isolated cDNA sequences are more conserved at the C-terminus compared to the N-terminal region which are more diverse. In addition, the alignment also determined a well conserved region with 14 aa residues which was previously reported to be specific to *VTE2* type amino acid sequences comprising HPT and HGGT (Venkatesh et al., 2006) (Fig. 1). Analysis of their domain structures further manifested that all the enzymes contained a relatively conserved prenyltransferase domain that span amino acid residues 175 to 455 and residues 110 to 393 for HGGT and HPT enzymes, respectively.

The presence of plastid targeting signal peptide was not observed in *EoHGGT* using SignalP and ChloroP programs. Nevertheless, WoLF PSORT results indicated that all the prenyltransferases successfully isolated in this study have been located in the plant plastid compartment. Results from protein domain searches indicated that all four isolated genes were classified under the UbiA prenyltransferase family. In addition, pairwise alignment showed that they shared high similarity in their 3′-terminus region. Therefore, all four genes were grouped in the subsequent phylogenetic analysis (Fig. 2). The phylogenetic tree clearly showed that the three HPT protein sequences of cyanobacteria were grouped into an out group from plant HGGT and HPT protein sequences. Therefore, it is suggested that prenyltransferase enzymes from cyanobacteria and plants may share a common ancestor and subsequently prenyltransferases have evolved independently in plants. Meanwhile all the plant prenyltransferases were clustered into two distinct subgroups, one for all HPTs and the other one for all HGGTs. This is most probably due to the different substrate preference for these two enzymes during the condensation step in vitamin E biosynthesis pathway (Yang et al., 2011). Orthology relationship was clearly observed in both HGGT and HPT specific subgroups with distinct clustering for monocots and dicots.

3.2. Gene expression analysis of oil palm HPT genes

Selection of stably expressed reference genes for each experiment condition is pivotal for obtaining a reliable gene expression result by RT-qPCR. In the present study, the expression stability of the three candidate reference genes (ACT, TUB and CYP) was evaluated by the geNorm program for all three experimental sets. These genes are commonly used as internal control genes in real-time quantitative studies of various plant species (Gao et al., 2004; Yeoh et al., 2012; Xu and Shi, 2006). However, a fluctuation in the expression of ACT gene was observed between two oil palm species in this work from the geNorm analysis. ACT gene was stably expressed in developing mesocarp of *E. guineensis* but its expression was inconsistent throughout the various developmental stages in *E. oleifera* mesocarp tissues. The results demonstrated the expression variation of a typical housekeeping gene in mesocarp of these two closely related oil palm species after examination. This further evidenced the importance of reference gene validation in order to avoid the misleading results. The geNorm software was used to select the most stably expressed reference genes in each set by calculating the average expression stability (M value) of all three genes. Two reference genes were identified for each set of samples using the recommended M values below the threshold of 1.5. For sets 1 and 3, ACT and CYP showed the highest expression stability while CYP and TUB showed the lowest M values in set 2.

Subsequently, the transcript profiles for oil palm HGGT and HPT were established by normalizing the expression of each gene of interest with the two identified stable control genes. At the earlier developmental stages of mesocarp, the expression profile of *EoHGGT* was not totally different from *EgHGGT* because both species showed reduced mRNA levels during 7, 10, 12 and 15 w.a.a (Fig. 3). However, the expression level of *EgHGGT* in *E. guineensis* mesocarp at 17 and 19 w.a.a demonstrated 17-fold and 54-fold of significant increments, respectively, compared to the expression at relatively younger mesocarps. Meanwhile the expression level of *EoHGGT* was highest at 17 w.a.a (5.8-fold). In contrast to oil palm *HGGT*, the expression of *EgHPT* showed a similar expression profile across developmental stages (Fig. 4). For *EoHPT*, a significant up-regulation was observed at 17 w.a.a with 21-fold of increment which declined at 19 w.a.a (Fig. 4). Moreover, the expression level of oil palm *HGGT* and *HPT* was compared in *E. guineensis* and *E. oleifera* 15 w.a.a kernel tissues, correspondingly. Interestingly, the expression level of *EoHGGT* at 15 w.a.a kernel tissues was significantly higher in

Fig. 1. Identification of a highly conserved region across oil palm HGGT and HPT amino acid sequences and their homologs using ClustalW alignment tool. (*) denotes single fully conserved residue; (:) denotes conservation of strong group; (.) denotes conservation of weak group. HPT sequences: KF974530, *EoHPT*; KF974529, *EgHPT*; ABB70124, *Allium porrum*; ABB70123, *Triticum aestivum*; ABB70122, *Zea mays*; ACN78585, *Lactuca sativa*; EU402477, *Sesamum indicum*; EU407262, *Angelica gigas*; ACC86836, *Coriandrum sativum*; AY089963, *Arabidopsis thaliana*; YP_007051923, *Nostoc sp.* PCC 7107; YP_007117098, *Oscillatoria nigro-viridis* PCC 7112; YP_007167807, *Halothece* sp. PCC 7418. HGGT sequences: EU402478, *Angelica gigas*; ADG26669, *Foeniculum vulgare*; ADG26671, *Carum carvi*; ADG26670, *Anethum graveolens*; ADG26668, *Daucus carota*; ADG26672, *Pimpinella anisum*; ADG26667, *Coriandrum sativum*; AB376093, *Hevea brasiliensis*; KP878511, *EgHGGT*; KP878512, *EoHGGT*; AY826350, *Vitis vinifera*; AY222861, *Triticum aestivum*; AY222860, *Hordeum vulgare*; AY222862, *Oryza sativa.* The boxes show the well conserved amino acid regions (14 aa residues) of all analyzed HGGT and HPT.

E. oleifera than *E. guineensis*. In contrast, the expression level of oil palm HPT was extremely lower in the kernel tissues as compared to oil palm HGGT. However, *EgHPT* was greatly up-regulated (56- and 16-fold) in spear leaves and young roots of *E. guineensis*. The transcript of *EgHGGT* in these two tissues was nearly undetected (Fig. 5).

3.3. Identification and in silico analysis of the EgHGGT and EgHPT promoter regions

In order to investigate the expression regulatory mechanisms of *EgHGGT* and *EgHPT* in *E. guineensis*, genome walking technique was employed to amplify both promoter regions and the resulting PCR fragments was sequenced. For *EgHPT*, the fragment contained 49 nucleotides of the 5′-UTR of *EgHPT* cDNA and the first base of the cDNA was designated +1 as the putative TSS. The TATA box and CAAT box which are classified as proximal regulatory elements were found at −54 bp and −220 of the predicted TSS, respectively. Meanwhile for the *EgHGGT* TSS, 'A' was detected 256 nucleotides upstream from the initial ATG. The TATA box and CAAT box were located at −29 bp and −124 bp upstream of the putative TSS, correspondingly. Relative to the TSS, the isolated *EgHPT* and *EgHGGT* promoter sequences were determined to be 951 bp and 1022 bp in length and designated as p*EgHPT* and p*EgHGGT*, respectively (Supplementary Figs. S1 and S2). Both promoter sequences were subjected to a promoter motif search on PLACE and PlantCARE databases. This led to the identification of a number of potential regulatory motifs corresponding to several known *cis*-acting elements of eukaryotic genes. Comparison of both promoter sequences revealed some common putative *cis*-acting regulatory motifs as listed in Table 3 together with their corresponding locations from the predicted TSS.

As shown in Table 3, a large number of light responsive elements (LREs) were detected in both promoter regions such as CIACADIANLELHC, GATA box, I-box core and SORLIP. GATA box is a common element present in many light-regulated promoters and recognized as an essential motif in expression of light dependent genes such as banana SPS gene (Jeong and Shih, 2003). Moreover, SORLIP

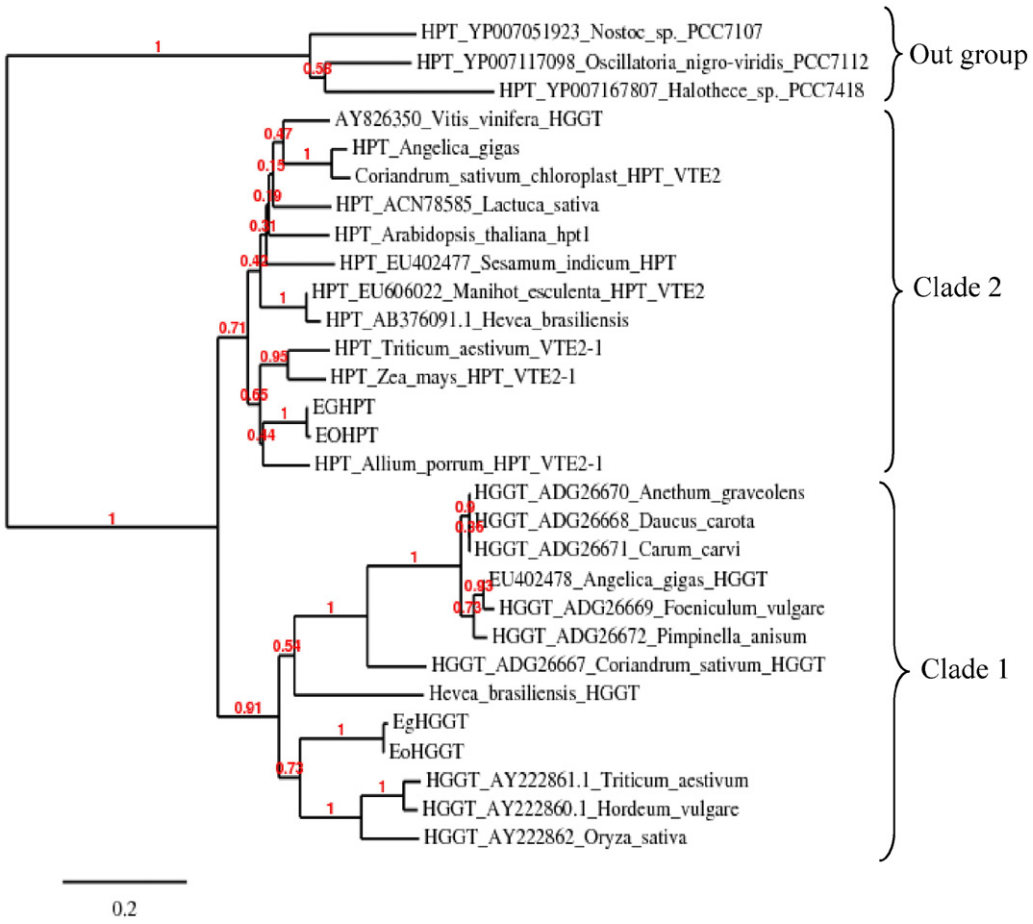

Fig. 2. Phylogenetic relationship among the derived amino acid sequences of the oil palm HGGT and HPT with other plants and cyanobacterias. All plant HGGTs and HPTs were grouped into subgroups 1 and 2, respectively. Three cyanobacteria HPTs were classified as an out-group. Bootstrap values are shown in red color. Scale bar shows genetic distance expressed as amino acid substitutions per site. The amino acid sequences are the same as in the multiple sequence alignment illustrated in Fig. 1.

was being reported to be over-represented in the phytochrome A-regulated promoters (Hudson and Quail, 2003). A single copy of the light *cis*-regulatory motif CIACADIANLELHC was found at positions −770 and −657 in the p*EgHPT* and p*EGHGGT*, respectively. This motif was identified as a conserved motif that appeared in tomato light-harvesting complex protein genes (*Lhc*) promoters which are expressed under the control of a circadian clock (Piechulla et al., 1998). Moreover, EE, a *cis*-element with the consensus sequence of AAAATATCT and known to be associated in circadian regulation of transcription was also detected at position −251 in the promoter of *EgHPT* as single copy (Harmer et al., 2000). This conserved motif was reported to be necessary for evening-specific transcription of *CAT3* promoter in *Arabidopsis* via deletion and site-mutagenesis analysis (Michael and McClung, 2002).

There are two regulatory motifs responsive to the dehydration stress, WAACCA and CANNTG motifs found in the 5′ upstream region

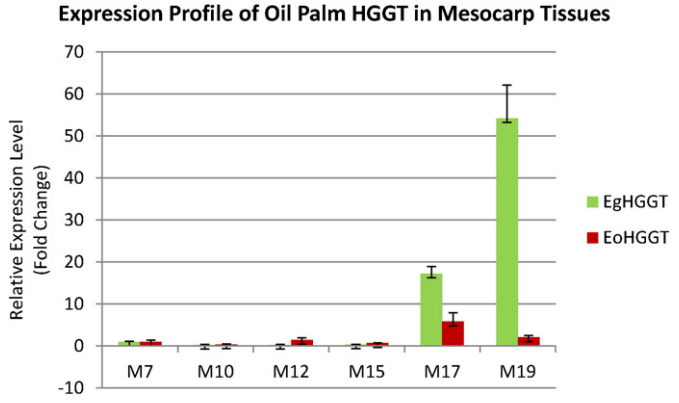

Fig. 3. Relative abundances of *EgHGGT* and *EoHGGT* in mesocarp tissues at six different developmental stages (M7–M19) of *E. guineensis* and *E. oleifera*, respectively. The value of the calibrator (M7) was set at 1 and the data of other tissues was then normalized to this value. Each gene was analyzed using one biological replicate with three technical replicates. Error bars indicated the standard error for three technical replicates.

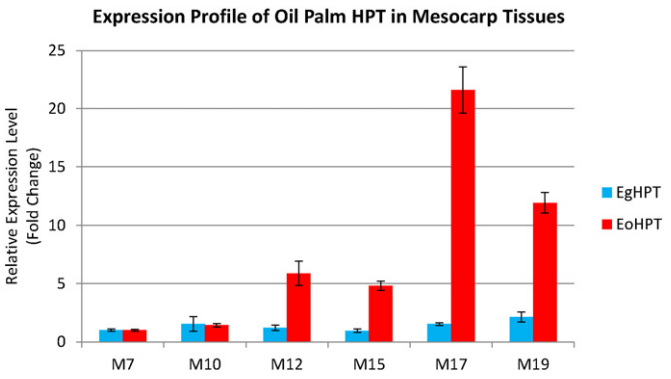

Fig. 4. Relative abundances of *EgHPT* and *EoHPT* in mesocarp tissues at six different developmental stages (M7–M19) of *E. guineensis* and *E. oleifera*, respectively. The value of the calibrator (M7) was set at 1 and the data of other tissues was then normalized to this value. Each gene was analyzed using one biological replicate with three technical replicates. Error bars indicated the standard error for three technical replicates.

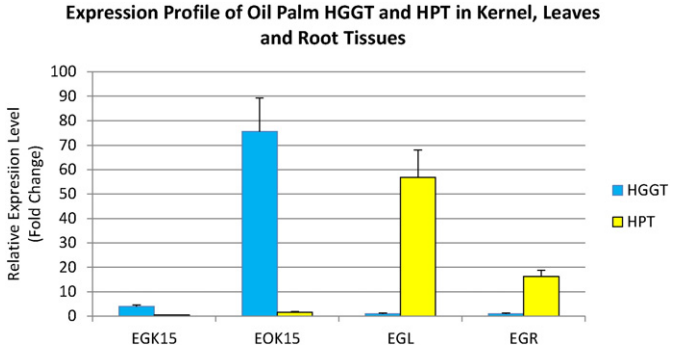

Fig. 5. Relative abundances of *EgHGGT* in *E. guineensis* 15 w.a.a kernel (EGK15), spear leaves (EGL) and young root (EGR); *EoHGGT* in *E. oleifera* 15 w.a.a kernel (EOK15); *EgHPT* in EGK15, EGL and EGR and *EoHPT* in EOK15. In this experiment, expression level of *EgHGGT* in EGL was the lowest and thus selected as a calibrator. The value of the calibrator was set at 1 and the data of other tissues was then normalized to this value. Each gene was analyzed using one biological replicate with three technical replicates. Error bars indicated the standard error for three technical replicates.

of p*EgHGGT*. These motifs have been demonstrated to bind with the corresponding MYC and MYB transcription factors and cooperatively activated the expression of a dehydration responsive RD22 gene in *Arabidopsis* (Abe et al., 1997; Neves-Borges et al., 2012). The presence of several putative ABRE was discovered in the 5′ upstream region of *EgHPT*. In plants, dehydration stress appears to be mediated via two signaling pathways; one is ABA-dependent and the other ABA-independent. ABA-responsive element (ABRE) with the core sequence PyACGTG/TC is a prominent *cis*-acting element required for ABA-dependent gene expression whereas dehydration-responsive element (DRE) is playing a key role in the ABA-independent pathway (Kim et al., 2011). Numerous ABREs have been identified and well-studied through functional dissection of the promoter of ABA-induced genes (Busk and Pages, 1998). In order to confer the ABA-responsive transcription, multiple ABREs or combinations of ABREs with coupling elements are needed (Narusaka et al., 2003; Go'mez-Porras et al., 2007). DRE-core motif (RCCGAC) was observed at position −910. This element has been demonstrated to function as a dehydration-responsive element (DRE) and also as a coupling element of ABRE in response to ABA through base-substitution analysis in transgenic tobacco plants (Narusaka et al., 2003). DRE was first identified as a fundamental *cis*-acting element in the *Arabidopsis* rd29A promoter for ABA-independent expression in response to dehydration and cold stresses (Yamaguchi-Shinozaki and Shinozaki, 1994).

Both promoter regions contained two GA-responsive elements, pyrimidine box and TAACAAA box. A similar consensus sequence (TAAC AAA/G) was identified as a core motif for GA responsiveness of α-amylase gene in barley aleurone cells (Skriver et al., 1991). The uniqueness of this GA-responsive element has been well defined in several studies using site-directed mutagenesis method and recognized as a gibberellin response element (GARE) since it plays the central role in GA responsiveness (Rogers and Rogers, 1992). Remarkably, an extraordinary TATCCAC box was observed in *EgHGGT* promoter region but not in 5′-upstream region of *EgHPT*.

4. Discussion

The identity of the common protein domain verified the role of the *E. guineensis* and *E. oleifera* HGGT and HPT in the vitamin E biosynthesis pathway which are catalyzing the condensation of HGA and a prenyldiphosphate (Hunter and Cahoon, 2007). Computational analysis using WoLF PSORT suggested that these oil palm prenyltransferases are plastid localized. Biosynthesis of tocopherol occurs in the plastids. Based on structural similarity and subcellular localization study, tocotrienols are also predicted to be synthesized in plastids (Schultz, 1990; Arango and Heise, 1997; Munné-Bosch and Alegre, 2002; Yang et al., 2011).

Both oil palm HGGT and HPT enzymes were grouped into the monocot-specific subgroup. This indicated that the phylogeny of HGGT and HPT in both monoecious palms is more closely related to the other monocots compared to dicots. A similar evolutionary structure was also demonstrated in the phylogenetic study performed using rice vitamin E biosynthesis genes and other reported tocochromanol biosynthesis genes (Chaudhary and Khurana, 2009). It is believed that the prenyltransferase genes have undergone many changes in the duplication process that contribute to the substrate specificity and spatial expression pattern.

In oil palm, the transcription of HGGT is strongly correlated to the oil deposition period in mesocarp tissues. Oil deposition in the oil palm mesocarp starts at around 15 w.a.a and continues until fruit maturity at about 20 w.a.a (Hartley, 1988). Our results indicate that the expression level of oil palm HGGT is greatly up-regulated in 17 and 19 w.a.a mesocarp compared with the younger fruit ages. Consistent with the current observation, accumulation of tocotrienols in *E. guineensis* mesocarp was only being detected after 16 w.a.a as described by Choo et al. (2004). In addition to mesocarp tissues, the transcript of oil palm HGGT was detected at high level in 15 w.a.a kernel tissues but almost negligible in leaves and roots. At 15 w.a.a, the oil deposition period is almost completed in kernel tissues (Hartley, 1988). Similarly, Cahoon et al. (2003) reported that barley HGGT mRNA was not detectable in leaves and roots of barley but found specifically in seeds. Moreover, a study conducted by Horvath et al. (2006) concluded that tocotrienols are highly associated with fruits and seeds rather than leaves. Seed dissection of

Table 3
List of putative *cis*-regulatory elements found in p*EgHPT* and p*EGHGGT* promoters where (+) is calculated from the positive strand and (−) is calculated from the negative strand based on the location of the TSS.

Elements	*Cis*-element	Sequences	Position from TSS in p*EgHPT*	Position from TSS in p*EgHGGT*
Gibberellins response	Pyrimidine box	CCTTTT	44(+), 874(+)	19 (−)
	GAREAT	TAACAAR	924(+)	654(+)
Copper response	CURECORECR	GTAC	578(+), 578(−)	53(+), 448(+), 53(−), 448(−)
Light regulation	CIACADIANLELHC	CAANNNNATC	770(+)	657(+)
	GATA box	GATA	115(+), 169(+), 218(+), 262(+), 19(−), 130(−), 188(−), 255(−), 389(−), 409(−), 776(−), 835(−)	365(+), 395(+), 419(+), 652(+), 708(+), 872(+), 882(+), 929(+), 1048(+), 1079(+), 236(−), 506(−), 580(−), 884(−), 903(−), 908(−), 954(−), 982(−)
	IBOXCORE	GATAA	218(+)	652(+), 872(+), 1079(+), 505(−), 579(−)
	SORLIP	GCCAC	302(−), 546(−), 668(−), 722(−)	16(−)
	GT1CONSENSUS	GRWAAW	183(+), 218(+), 318(+), 540(+), 942(+), 876(−)	543(+), 566(+), 748(+), 844(+), 872(+), 939(+), 1079(+), 504(−), 525(−), 607(−), 200(−), 215(−), 536(−), 578(−)
	Box 4	ATTAAT	103(+)	298(+)
Dehydration response	MYBCORE	CNGTTR	13(+), 854(+), 459(−), 689(−), 770(−)	154(+), 151(−)
	MYCCONSENSUS	CANNTG	7(+), 13(+), 35(+), 501(+), 675(+), 7(−), 13(−), 35(−), 501(−), 675(−)	48(+), 220(+), 308(+), 855(+), 48(−), 220(−), 308(−), 855(−)
Heat stress	HSE	AAAAAATTTC	887(−)	688(+), 719(+)

grape also demonstrated that tocotrienols uniquely occurred in the endosperm tissues of grape (Horvath et al., 2006). Detection of HGGT transcripts in oil-bearing tissues, in this case referred to as mesocarp and kernel of oil palm but not in other tissues suggests that tocotrienols might principally function in protecting storage oil from oxidative damage and in promoting the longevity of monocot seeds as it is well known as powerful lipid soluble antioxidant agents (Yang et al., 2011). Meanwhile, the oil palm HPT was found to be expressed in all the tissues except kernel and this is in agreement with the presence of tocopherols in leaves and mesocarp but absence in kernel as reported by Kato et al. (2002). The high expression of *EgHPT* in leaves probably correlates with potential accumulation of tocopherol in this organ, since tocopherols are well recognized for their ability to protect photosynthetic tissues from photooxidation and photoinactivation (Collakova and DellaPenna, 2003; Havaux et al., 2005). These results may suggest that the expression of these genes is spatially and temporally regulated.

Despite the HGGT and HPT displaying different transcription profiles in various developmental stages of the mesocarp in *E. guineensis* and *E. oleifera*, vitamin E composition in the individual oil is comparable with each other. The major vitamin E isomers present in both palm oils are α-tocotrienol (29%), α-tocopherol (28%), γ-tocotrienol (28%) and δ-tocotrienol (14%). However, palm oil derived from *E. oleifera* exhibited higher vitamin E content (700–1500 ppm) as compared with *E. guineensis* (600–1000 ppm) (Jalani et al., 1997). In the past few years, several genetic engineering strategies have been developed to fortify the vitamin E content in oilseed crops. Overexpressing a single HPT transgene under the regulation of a seed-specific promoter in *A. thaliana* and soybean only yield modest increase of tocopherols content (less than 1.5 fold) (Savidge et al., 2002; Karunanandaa et al., 2005). Whereas seed-specific expression of HGGT transgene in corn seeds showed up to six-fold increment of tocochromanols content, mostly in the form of tocotrienols (Cahoon et al., 2003). Therefore, up-regulating the oil palm HGGT expression could potentially be an efficient strategy to enhance the total vitamin E content in palm oil especially for commercial production of tocotrienols due to the expanding needs in nutraceuticals market.

The occurrence of numerous LREs in the promoters of both genes might indicate that the expression of *EgHPT* and *EgHGGT* is associated with response to light. This is consistent to the previous studies showing that the level of tocopherol in plant tissue is related to the light intensity (Lichtenthaler, 2007). In the study, the total tocopherol levels increased exponentially in *Arabidopsis* after subjected to high light intensity. The results correlated to the significantly elevated HPT mRNA level and HPT specific activity (Collakova and DellaPenna, 2003).

Under natural environment, plants are exposed to assorted types of abiotic factors such as stresses caused by drought, cold, and high salinity. There were two regulatory motifs responsive to the dehydration stress, WAACCA and CANNTG, found in the 5′ upstream region of *EgHGGT*. The presence of several putative ABRE and MYB core motifs was also discovered in the 5′ upstream region of *EgHPT*. This result was in agreement with a promoter analysis done on all tocopherol biosynthetic genes in tomato which also detected several significantly overrepresented MYB transcription factor binding sites and ABA-responsive motifs (Quadrana et al., 2013). The authors suggested that MYB transcription factors might be the main regulatory proteins involved in the tocopherol biosynthesis pathway as the identified motifs possess a conserved positional preference in the first 200 bp of the promoters when mapped on the tocopherol pathway genes. Tocochromanols act as the strong antioxidants by protecting the plant membranes from lipid peroxidation, thus their formation and degradation were predicted to be influenced by the degree of stress (Andersson et al., 2008). The effect of these abiotic factors on the formation of tocopherols has been reviewed by Munné-Bosch and Alegre (2002) and concluded that the synthesis rate of tocopherols could be induced by a low level of stress while the high level of stress will overwhelm the capacity of antioxidant system and results in lower content of tocopherols. Over the years, increasing the level of α-tocopherol under drought stress was being reported in numerous plant species including *Arabidopsis*, tobacco and rosemary (Munné-Bosch et al., 1999; Liu et al., 2008).

Phytohormones are known for their essential roles in plant development through regulating the expression of the related genes. Promoter regions of both oil palm genes contained two GA-responsive elements, pyrimidine box and TAACAAA box. Gibberellins (GAs) as the outstanding phytohormones associated with various aspects of the plant growth including seed germination, leaf expansion, stem elongation, flowering, and fruit development (Ogawa et al., 2003). Earlier reports suggested that GA plays an important role in overcoming seed dormancy and stimulating seed germination in *Arabidopsis* (Debeaujon and Koornneef, 2000). TATCCAC box was observed in *EgHGGT* promoter region but not in 5′-upstream region of *EgHPT*. The combination of TAACAAA box, TATCCAC box and pyrimidine box was identified as a gibberellin responsive complex (GARC) and act cooperatively to give a high level of GA-regulated expression (Gubler and Jacobsen, 1992; Rogers et al., 1994).

5. Conclusion

In this study, we have identified the HGGT and HPT from two important oil palm species that appear to contain the UbiA prenyltransferase conserved domain and are more evolutionarily related to their counterparts from the other monocot plant species. Quantitative gene expression analysis demonstrated that HPT was constitutively expressed in the vegetative and fruit tissues. However, the HGGT showed preferential expression in the oil-bearing mesocarp and kernel tissues and was highly expressed at the fruit ripening stage of the mesocarp supporting its role in providing protection to the vegetable oils from oxidative damage. A number of important *cis*-regulatory elements related to phytohormone-responsive, light-responsive and abiotic factor-responsive which may be involved in coordinating expression of both genes were identified. The information generated could pave the way for improvement of oil palm vitamin E content through genetic engineering and development of molecular markers to support intraspecific and interspecific breeding approaches.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.plgene.2016.01.003.

Acknowledgment

The authors thank the School of Graduate Studies of Universiti Putra Malaysia for providing a Graduate Research Fellowship to the first author.

References

Abe, H., Yamaguchi-Shinozaki, K., Urao, T., Iwasaki, T., Hosokawa, D., Shinozaki, K., 1997. Role of *Arabidopsis* MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. Plant Cell 9, 1859–1868.

Andersson, S.C., Rumpunen, K., Johhansson, E., Olsson, M.E., 2008. Tocopherols and tocotrienols in Sea Buckthorn (*Hippophae rhamnoides* L.) berries during ripening. J. Agric. Food Chem. 56, 6701–6706.

Arango, Y., Heise, K.P., 1997. α-Tocopherol synthesis by *Capsicum* fruit chromoplasts. J. Plant Physiol. 150, 509–513.

Bramley, P.M., Elmadfa, I., Kafatos, A., Kelly, F.J., Manios, Y., Roxborough, H.E., Schuch, W., Sheehy, P.J.A., Wagner, K.H., 2000. Review—vitamin E. J. Sci. Food Agric. 80, 913–938.

Busk, P.K., Pages, M., 1998. Regulation of abscisic acid-induced transcription. Plant Mol. Biol. 37 (3), 425–435.

Cahoon, E.B., Hall, S.E., Ripp, K.G., Ganzke, T.S., Hitz, W.D., Coughlan, S.J., 2003. Metabolic redesign of vitamin E biosynthesis in plants for tocotrienol production and increased antioxidant content. Nat. Biotechnol. 21 (9), 1082–1087.

Chaudhary, N., Khurana, P., 2009. Vitamin E biosynthesis genes in rice: molecular characterization, expression profiling and comparative phylogenetic analysis. Plant Sci. 177, 479–491.

Choo, Y.M., Ma, A.N., Chuah, C.H., Khor, H.T., Bong, S.C., 2004. A developmental study on the appearance of tocopherols and tocotrienols in developing palm mesocarp (*Elaeis guineensis*). Lipids 39 (6), 561–564.

Collakova, E., DellaPenna, D., 2001. Isolation and functional analysis of homogentisate phytyltransferase from *Synechocystis* sp. PCC 6803 and *Arabidopsis*. Plant Physiol. 127, 1113–1124.
Collakova, E., DellaPenna, D., 2003. Homogentisate phytyltransferase activity is limiting for tocopherol biosynthesis in *Arabidopsis*. Plant Physiol. 13, 632–642.
Debeaujon, I., Koornneef, M., 2000. Gibberellin requirement for *Arabidopsis* seed germination is determined both by testa characteristics and embryonic abscisic acid. Plant Physiol. 122 (2), 415–424.
Dörmann, P., 2003. Corn with enhanced antioxidant potential. Nat. Biotechnol. 21 (9), 1015–1016.
Doyle, J.J., Doyle, J.L., 1990. Isolation of plant DNA from fresh tissues. Focus 12, 13–15.
Gao, X., Jackson, T.A., Lambert, K.N., Li, S., 2004. Detection and quantification of *Fusarium solani* f. sp. *glycines* in soybean roots with real-time quantitative polymerase chain reaction. Plant Dis. 88 (12), 1372–1380.
Go'mez-Porras, J.L., Riaño-Pachón, D.M., Dreyer, I., Mayer, J.E., Mueller-Roeber, B., 2007. Genome-wide analysis of ABA-responsive elements ABRE and CE3 reveals divergent patterns in *Arabidopsis* and rice. BMC Genomics 8, 260.
Gubler, F., Jacobsen, J.V., 1992. Gibberellin-responsive elements in the promoter of a barley high-pl α-amylase gene. Plant Cell 4, 1435–1441.
Harmer, S.L., Hogenesch, J.B., Straume, M., Chang, H.S., Han, B., Zhu, T., Wang, X., Kreps, J.A., Kay, S.A., 2000. Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. Science 290, 2110–2113.
Hartley, C.W.S., 1988. The Oil Palm. third ed. Longman, New York.
Havaux, M., Eymery, F., Porfirova, S., Rey, P., Dörmann, P., 2005. Tocochromanol protects against photo inhibition and oxidative stress in *Arabidopsis thaliana*. Plant Cell 17, 3451–3469.
Horvath, G., Wessjohann, L., Bigirimana, J., Jansen, M., Guisez, Y., Caubergs, R., Horemans, N., 2006. Differential distribution of tocopherols and tocotrienols in photosynthetic and non-photosynthetic tissues. Phytochemistry 67, 1185–1195.
Hudson, M.E., Quail, P.H., 2003. Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. Plant Physiol. 133, 1605-161.
Hunter, S.C., Cahoon, E.B., 2007. Enhancing vitamin E in oilseeds: unraveling tocopherol and tocotrienol biosynthesis. Lipids 42, 97–108.
Jalani, B.S., Cheah, S.C., Rajanaidu, N., Darus, A., 1997. Improvement of oil palm through breeding and biotechnology. JAOCS 74, 1451–1455.
Jeong, M.J., Shih, M.C., 2003. Interaction of a GATA factor with *cis*-acting elements involved in light regulation of nuclear genes encoding chloroplast glyceraldehyde-3-phosphate dehydrogenase in *Arabidopsis*. Biochem. Biophys. Res. Commun. 300, 555–562.
Kamal-Eldin, A., Appelqvist, L.A., 1996. The chemistry and antioxidant properties of tocopherols and tocotrienols. Lipids 31 (7), 671–701.
Karunanandaa, B., Qi, Q., Hao, M., Baszis, S.R., Jensen, P.K., Wong, Y.H., Jiang, J., Venkatramesh, M., Gruys, K.J., Moshiri, F., Post-Beittenmiller, D., Weiss, J.D., Valentin, H.E., 2005. Metabolically engineered oilseed crops with enhanced seed tocopherol. Metab. Eng. 7, 384–400.
Kato, A., Yamaoka, M., Gapor, A., Berger, K.G., 2002. Tocopherols of oil palm leaf. JAOCS 60 (12). http://dx.doi.org/10.1007/BF02669973.
Kim, J.S., Mizoi, J., Yoshida, T., Fujita, Y., Nakajima, J., Ohori, T., Todaka, D., Nakashima, K., Hirayama, T., Shinozaki, K., Yamaguchi-Shinozaki, K., 2011. An ABRE promoter sequence is involved in osmotic stress-responsive expression of the DREB2A gene, which encodes a transcription factor regulating drought-inducible genes in *Arabidopsis*. Plant Cell Physiol. 52 (12), 2136–2146.
Lichtenthaler, H.K., 2007. Biosynthesis, accumulation and emission of carotenoids, α-tocopherol, plastoquinone and isoprene in leaves under high photosynthetic irradiance. Photosynth. Res. 92, 163–179.
Liu, X., Hua, X., Guo, J., Qi, D., Wang, D., Liu, Z., Jin, Z., Chen, S., Liu, G., 2008. Enhanced tolerance to drought stress in transgenic tobacco plants overexpressing VTE1 for increased tocopherol production from *Arabidopsis thaliana*. Biotechnol. Lett. 30, 1275–1280.
Michael, T.P., McClung, C.R., 2002. Phase-specific circadian clock regulatory elements in *Arabidopsis*. Plant Physiol. 130, 627–638.
Munné-Bosch, S., Alegre, L., 2002. The function of tocopherols and tocotrienols in plants. Crit. Rev. Plant Sci. 21, 31–57.
Munné-Bosch, S., Schwarz, K., Alegre, L., 1999. Enhanced formation of α-tocopherol and highly oxidized abietane diterpenes in water-stressed rosemary plants. Plant Physiol. 121, 1047–1052.
Narusaka, Y., Nakashima, K., Shinwari, Z.K., Sakuma, Y., Furihata, T., Abe, H., Narusaka, M., Shinozaki, K., Yamaguchi-Shinozaki, K., 2003. Interaction between two *cis*-acting elements, ABRE and DRE, in ABA-dependent expression of *Arabidopsis* rd29A gene in response to dehydration and high-salinity stresses. Plant J. 34, 137–148.
Neves-Borges, A.C., Guimarães-Dias, F., Cruz, F., Mesquita, R.O., Nepomuceno, A.L., Romano, E., Loureiro, M.E., de Fátima Grossi-de-Sá, M., Alves-Ferreira, M., 2012. Expression pattern of drought stress marker genes in soybean roots under two water deficit systems. Genet. Mol. Biol. 35 (1), 212–221.
Ogawa, M., Hanada, A., Yamauchi, Y., Kuwahara, A., Kamiya, Y., Yamaguchi, S., 2003. Gibberellin biosynthesis and response during *Arabidopsis* seed germination. Plant Cell 15, 1591–1604.
Piechulla, B., Merforth, N., Rudolph, B., 1998. Identification of tomato *Lhc* promoter regions necessary for circadian expression. Plant Mol. Biol. 38, 655–662.
Prescott, A., Martin, L., 1987. A rapid method for quantitative assessment of levels of specific mRNAs in plants. Plant Mol. Biol. Report. 4, 219–224.
Quadrana, L., Almeida, J., Otaiza, S.N., Duffy, T., Corrêa da Silva, J.V., de Godoy, F., Asís, R., Bermúdez, L., Fernie, A.R., Carrari, F., Rossi, M., 2013. Transcriptional regulation of tocopherol biosynthesis in tomato. Plant Mol. Biol. 81, 309–325.
Rasid, O.A., Wan Nur Syuhada, W.S., Nor Hanin, A., Masura, S.S., Zulqarnain, M., Ho, C.L., Sambanthamurthi, R., Suhaimi, N., 2008. RT-PCR amplification and cloning of partial DNA sequence coding for oil palm (*Elaeis oleifera*) phytoene synthase gene. Asia-Pac. J. Mol. Biol. 16 (1), 17–24.
Rival, A., Jaligot, E., Beule, T., Jean Finnegan, E., 2008. Isolation and expression analysis of genes encoding *MET*, *CMT*, and *DRM* methyltransferases in oil palm (*Elaeis guineensis* Jacq.) in relation to the 'mantled' somaclonal variation. J. Exp. Bot. 59 (12), 3271–3281.
Rogers, J.C., Rogers, S.W., 1992. Definition and functional implications of gibberellin and abscisic acid *cis*-acting hormone response complexes. Plant Cell 4, 1443–1451.
Rogers, J.C., Lanahan, M.B., Rogers, S.W., 1994. The *cis*-acting gibberellin response complex in high-pI-amylase gene promoters. Plant Physiol. 105, 151–158.
Sambrook, J., Fritsch, E.F., Maniatis, T., 2001. Molecular Cloning: A Laboratory Manual. third ed. Cold Spring Harbor Laboratory Press, New York.
Savidge, B., Weiss, J.D., Wong, Y.H., Lassner, M.W., Mitsky, T.A., Shewmaker, C.K., Post-Beittenmiller, D., Valentin, H.E., 2002. Isolation and characterization of homogentisate phytyltransferase genes from *Synechocystis* sp. PCC 6803 and *Arabidopsis*. Plant Physiol. 129, 321–332.
Schultz, G., 1990. Biosynthesis of α-tocopherols in chloroplasts of higher plants. Eur. J. Lipid Sci. Technol. 92, 86–91.
Sen, C.K., Khanna, S., Roy, S., 2006. Tocotrienols: vitamin E beyond tocopherols. Life Sci. 78, 2088–2098.
Skriver, K., Olsen, F.L., Rogers, J.C., Mundy, J., 1991. *Cis*-acting DNA elements responsive to gibberellin and its antagonist abscisic acid. Proc. Natl. Acad. Sci. U. S. A. 88, 7266–7270.
Vandesompele, J., Preter, K.D., Pattyn, F., Poppe, B., Roy, N.V., Paepe, A.D., Speleman, F., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 3 (7), 34.1–34.11.
Venkatesh, T.V., Karunanandaa, B., Free, D.L., Rottnek, J.M., Baszis, S.R., Valentin, H.E., 2006. Identification and characterization of an *Arabidopsis* homogentisate phytyltransferase paralog. Planta 223, 1134–1144.
Xu, W.F., Shi, W.M., 2006. Expression profiling of the 14-3-3 gene family in response to salt stress and potassium and iron deficiencies in young tomato (*Solanum lycopersicum*) roots: analysis by real-time RT-PCR. Ann. Bot. Lond. 98, 965–974.
Yamaguchi-Shinozaki, K., Shinozaki, K., 1994. A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature or high-salt stress. Plant Cell 6, 251–264.
Yang, W.Y., Cahoon, R.E., Hunter, S.C., Zhang, C.Y., Han, J.X., Borgschulte, T., Cahoon, E.B., 2011. Vitamin E biosynthesis: functional characterization of the monocot homogentisate geranylgeranyl transferase. Plant J. 65, 206–217.
Yeoh, K.A., Othman, A., Meon, S., Abdullah, F., Ho, C.L., 2012. Sequence analysis and gene expression of putative exo- and endo-glucanases from oil palm (*Elaeis guineensis*) during fungal infection. J. Plant. Physiol. 169 (15), 1565–1570.